

# Reducing Complexity in Next-Generation Multi-User MIMO Systems

KONSTANTINOS NTOUGIAS    DIMITRIOS NTAIKOS    CONSTANTINOS B. PAPADIAS  
Athens Information Technology (AIT), 44 Kifissias Avenue, 15125 Maroussi, Greece.

kontou | dint | cpap@ait.gr

## Abstract

Recently, several advanced multi-antenna radio communications technologies have emerged to meet the increased capacity demands in wireless multi-user networks. Despite their great potential, the extent of these techniques' practical applicability still remains questionable, since they have to face either backhaul limitations or cost and hardware constraints. In this paper, we propose a new system solution which includes network architecture, antenna technology and radio transmission protocol to reduce drastically the hardware complexity and cost as well as the channel state information / user data feedback requirements of multi-user multi-antenna wireless networks. We focus on the forward link of an interference channel in a cloud radio access network setup wherein an arbitrary number of remote radio heads are each equipped with a single radio frequency module parasitic antenna array and wish to send data to their respective single-antenna user terminals, while co-existing in time and frequency. Base stations select cooperatively the optimal combination of pre-determined beams prior transmission. Our proposed approach is able to achieve the aforementioned goals, while offering significant downlink sum-rate gains due to the available spatial degrees of freedom.

**Index terms**— Remote Radio Head (RRH), Electronically Steerable Parasitic Array Radiator (ESPAR) Antenna, Channel State Information (CSI), Zero-Forcing (ZF) Precoding, Interference Channel (IFC).

## 1 Introduction

Over the last decade, an enormous growth of mobile data traffic has been witnessed. As several studies indicate [1][2], this trend is likely to continue in the years to come in an even more abrupt pace. In view of the well-known scarcity and high cost of the radio spectrum, current cellular mobile broadband networks incorporate multiple-input multiple-output (MIMO) technology as a response to the

increased capacity demands.

Traditional single-user MIMO (SU-MIMO) focuses on the physical layer performance of the communication between a base station (BS) and a mobile terminal (MT) over a point-to-point link. It exploits the additional degrees of freedom (DoFs) provided by the use of multiple antennas to enhance spectral efficiency through spatial multiplexing of individual data streams or to reduce co-channel interference by spatially focusing transmissions and separating co-channel signals.

Multi-user MIMO (MU-MIMO) is an evolution of SU-MIMO that enhances link-level performance by exploiting DoFs to spatially separate users, thus enabling a BS to communicate simultaneously with multiple mobile users (MUs) over a multipoint-to-point (uplink) or a point-to-multipoint (downlink) channel.

Recently, new MIMO paradigms, which aim at improving system-level performance, have been introduced. Their common characteristic is that they increase significantly the available DoFs to boost the performance of the system. More specifically, cooperative MIMO [3][4] enables the cooperation between different BSs to mitigate inter-cell interference and increase area spectral efficiency and system capacity, while massive MIMO [5][6] makes use of an excessive number of antennas at the cell site to orthogonalize MUs in the spatial domain.

It is expected that both these technologies will be an integral part of future fifth generation (5G) systems. However, their transition from theory into implementation requires the addressing of some important practical issues.

Cooperative MIMO refers to a collection of techniques with varying level of cooperation, from network MIMO which requires the sharing of user data and global channel state information (CSI) between the individual BSs to cooperative beamforming which requires only the exchange of precoding matrices between the cooperating nodes. In practice, backhaul latency limits the effectiveness of these techniques, while backhaul capacity constraints reduce the achievable level of cooperation

and the corresponding performance gains.

Massive MIMO, on the other hand, places a heavy burden, in terms of cost, to the mobile network operators due to the large number of radio-frequency (RF) chains (one per active antenna element) that is required for its implementation. Moreover, the tight spacing of antenna elements may lead to reduced efficiency due to spatial correlations.

In addition, as the number of BSs or/and antennas increases, channel estimation becomes a challenging task.

In this paper, we describe a system setup that is promising in its ability to address these challenges. We focus on the study of a radio transmission protocol that takes advantage of the attributes of the employed technologies to achieve substantial performance gains, while reducing significantly system complexity.

The structure of the remainder of the paper is as follows: In Section 2, the considered system architecture and antenna technology is presented. Next, the system and channel model is described in Section 3. We continue in Section 4 with a presentation of the proposed radio protocol. In Section 5 we present numerical performance results for different flavors of the proposed transmission scheme and compare them against performance bounds. We then discuss in Section 6 about the complexity reduction that is achieved with the considered radio protocol as well as about the relevant complexity-feedback overhead and complexity-performance tradeoffs. In Section 7 we study the performance of the transmission scheme when CSI is imperfect. Finally, in Section 8 we present our conclusions and discuss about future extensions of our work.

## 2 System Architecture and Technologies

According to the discussion in Section 1, we define three main criteria that are important to be met in order to enable the use of Cooperative / Massive MIMO in practice:

1. Low-cost implementation of systems with many antennas.
2. Low-delay communication between cooperating nodes.
3. Minimization of the information that has to be exchanged between the cooperating base stations.

In order to fulfill these requirements, all aspects of system design have to be taken into account, that is, network architecture, antenna technologies, and

radio transmission protocols. In this Section, we describe two proposed components that can meet the first two of these objectives. Next, we will focus on a proposed radio scheme that can be employed by this system to reduce feedback complexity.

### 2.1 ESPAR Antennas

Advanced parasitic antenna arrays, often called electronically steerable parasitic array radiators (ESPARs), provide multi-antenna functionality using fewer active elements (and, as a consequence, fewer RF units) than conventional antenna arrays. They accomplish this by adjusting the loads of parasitic elements that are placed in the vicinity of the active element(s), thus controlling the corresponding induced currents caused by mutual coupling to form and steer beams [7]. This concept is illustrated in Fig. 1.

ESPAR technology can be viewed as an attractive enabler of massive MIMO, since it allows the implementation of large antenna systems with much reduced complexity, size and cost.

### 2.2 Cloud-RAN/RRH

Cloud Radio Access Network (Cloud-RAN or C-RAN) [8][9] is a split base station architecture which separates the baseband unit (BBU) from the radio unit. More specifically, in C-RAN, remote radio units (RRUs), also called remote radio heads (RRHs), remain at the cell site, while BBUs are centralized and virtualized using cloud technology. Baseband and radio resources are allocated to “virtual base stations” in real time according to processing and radio coverage needs. Remote radio heads are connected with the centralized BBU pool through optical fibers. Protocols such as Common Public Radio Interface (CPRI) [10] are used to enable the communication between the virtualized baseband pool and the RRHs over this new network segment, which is often referred to as the mobile fronthaul.

C-RAN is hence considered to be a compelling candidate technology for bringing cooperative MIMO communication into reality, due to the centralization and virtualization of the baseband and the use of optical transmission technologies at the fronthaul which have the potential to meet the stringent delay and bandwidth requirements of these advanced techniques [11].

## 3 System and Channel Model

In this Section, we introduce the system and channel model of our setup. Before getting in the details, let us present the notation that we follow throughout the paper.

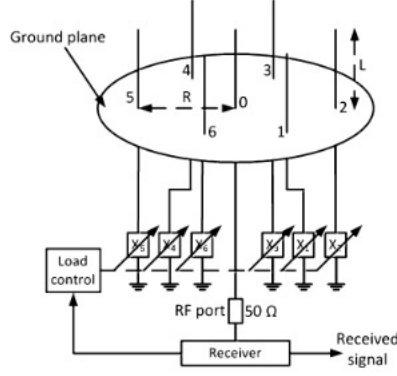


Figure 1: The ESPAR paradigm for beam shaping in the analog domain with digital control [7].

*Notation:* With  $a$ ,  $\mathbf{a}$ , and  $\mathbf{A}$  we denote a (in general, complex-valued) scalar, vector, and matrix, respectively.  $\mathbf{A}_{i,j}$  represents the  $(i,j)$  (row, column) element of  $\mathbf{A}$ . The Hermitian (transpose conjugate), determinant, and rank of  $\mathbf{A}$  are denoted by  $\mathbf{A}^\dagger$ ,  $\det(\mathbf{A})$  and  $\text{rank}(\mathbf{A})$ , respectively, while the inverse and the Moore-Penrose pseudo-inverse of  $\mathbf{A}$  are denoted by  $\mathbf{A}^{-1}$  and  $\mathbf{A}^+$ , respectively.  $\text{Tr}(\mathbf{A})$  represents the trace of  $\mathbf{A}$  and  $\|\mathbf{A}\|_F$  denotes its Frobenius norm, whereas  $\text{diag}(\mathbf{a})$  represents a diagonal matrix with the elements of vector  $\mathbf{a}$  at its main diagonal.  $\|\mathbf{a}\|$  stands for the Euclidean norm of vector  $\mathbf{a}$ .  $\mathbf{0}_n$  and  $\mathbf{I}_n$  denote the  $n \times n$  zero and identity matrix, respectively (the subscript indicating the dimension of the matrix may be omitted whenever is irrelevant).  $\mathbb{C}$  denotes the set of complex numbers,  $\mathbb{E}\{\cdot\}$  represents the expectation operator,  $X \sim \mathcal{CN}(0, \sigma^2)$  refers to a complex-valued random variable (RV)  $X$  following the Gaussian distribution with zero mean (i.e.,  $\mathbb{E}\{X\} = 0$ ) and variance  $\sigma_x^2$ , and  $\mathbf{a} \sim \mathcal{CN}(\mathbf{0}, \sigma_a^2 \mathbf{I})$  represents a complex-valued Gaussian vector with mean matrix  $\mathbb{E}\{\mathbf{a}\} = \mathbf{0}$  and covariance matrix  $\mathbf{R}_{aa} = \mathbb{E}\{\mathbf{a}\mathbf{a}^\dagger\} = \sigma_a^2 \mathbf{I}$ , i.e., the elements of  $\mathbf{a}$  are complex-valued RVs with zero mean and variance  $\sigma_a^2$ .

### 3.1 System Model

We consider the downlink of a C-RAN system with  $K$  RRHs, each representing an individual BS and being equipped with a single-RF ESPAR antenna, where each one of them wishes to communicate with a user equipment (UE) having a single antenna. All transmissions take place simultaneously and over the same frequency band.

Note that due to the C-RAN architecture, the BSs can efficiently share information. The MUs, on the other hand, do not cooperate.

Each ESPAR antenna is able to generate  $L$  distinct predetermined beams. According to some criterion, the best  $K$ -tuple of beams is selected for transmission from the  $L^K$  different beam combina-

tions in total, with each one of these beams being generated at a different RRH, as seen in Fig. 2. (The beam selection criteria are presented in Section 4.)

The corresponding “beam-domain transformed” channel between these  $K$  transmitter-receiver (TX-RX) pairs is equivalent to a single-input single-output (SISO)  $K$ -user interference channel (IFC), since each TX has a single active antenna. However, we should not ignore the fact that the TXs are equipped with parasitic antenna arrays and make use of beamforming, in contrast to conventional SISO TXs which utilize omni-directional antennas.

Mathematically, this SISO IFC with  $K$  TXs and  $K$  RXs having  $\{N_{T,k}\}_{k=1}^K = 1$  and  $\{N_{R,k}\}_{k=1}^K = 1$  active antennas each, respectively, can be viewed as a multi-user multiple-input single-output (MUMISO)  $K$ -user IFC formed by a composite TX with  $N_T = \sum_{k=1}^K N_{T,k} = K$  antennas and  $K$  RXs, each with  $\{N_{R,k}\}_{k=1}^K = 1$  antenna.

We assume narrowband transmission (i.e., flat-fading channel), such that no inter-symbol interference is caused by multipath. This is the case, for example, in indoor environments with relatively small delay spread as well as when orthogonal frequency division multiplexing (OFDM) transmission is employed to convert a frequency-selective channel into parallel frequency-flat sub-channels. We also assume block-fading, such that the channel remains fixed during the transmission of each symbol. Therefore, we can omit time dependence in the system model.

Under these assumptions, the baseband received signal at user  $k$ , where  $k = 1, \dots, K$ , for a selected beam combination, can be expressed as

$$y_k = \sum_{m=1}^K h_{k,m} s_m + n_k, \quad k = 1, \dots, K \quad (1)$$

where  $s_m \in \mathbb{C}$  is the complex symbol transmitted by  $\text{Tx}_m$ ,  $y_k \in \mathbb{C}$  is the complex received symbol at  $\text{Rx}_k$ ,  $n_k \sim \mathcal{CN}(0, \sigma_n^2)$  is complex additive Gaussian noise with zero mean and variance  $\sigma_n^2$ , and  $h_{k,m}$  is

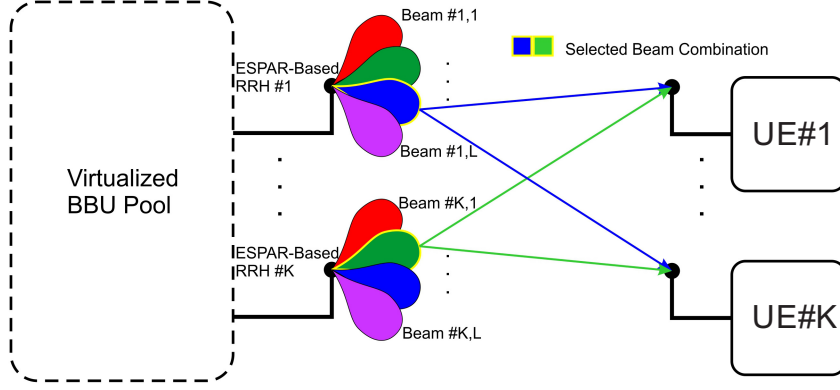


Figure 2: System model. Note the transformation of the channel to a “beam-domain” interference / broadcast channel.

the complex channel coefficient from  $\text{Tx}_m$  to  $\text{Rx}_k$ .

Equivalently, we can write

$$y_k = \mathbf{h}_k^\dagger \mathbf{s} + n_k, \quad k = 1, \dots, K \quad (2)$$

where  $\mathbf{h}_k \in \mathbb{C}^{K \times 1}$  is the vector of the channel coefficients between  $\text{Rx}_k$  and  $\{\text{Tx}_m\}_{m=1}^K$  and  $\mathbf{s} \in \mathbb{C}^{K \times 1}$  is the vector of transmitted symbols.

The input-output model of the overall channel is

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \quad (3)$$

where  $\mathbf{y} \in \mathbb{C}^{K \times 1}$  is the vector of received symbols;  $\mathbf{n} \in \mathbb{C}^{K \times 1}$  is a complex-valued Gaussian noise vector with  $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_K)$ , i.e., with mean matrix  $\mathbb{E}\{\mathbf{n}\} = \mathbf{0}_K$  and covariance matrix  $\mathbf{R}_{nn} = \mathbb{E}\{\mathbf{n}\mathbf{n}^\dagger\} = \sigma_n^2 \mathbf{I}_K$ ; and  $\mathbf{H} \in \mathbb{C}^{K \times K}$  is the channel matrix expressed as

$$\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K]^\dagger. \quad (4)$$

The total transmitted power is constrained to  $P$ , i.e.,

$$\text{Tr}(\mathbf{R}_{ss}) = \text{Tr}(\mathbb{E}\{\mathbf{s}\mathbf{s}^\dagger\}) \leq P, \quad (5)$$

where  $\mathbf{R}_{ss} = \mathbb{E}\{\mathbf{s}\mathbf{s}^\dagger\}$  is the covariance matrix of the transmitted signal vector  $\mathbf{s}$ .

For convenience, and without loss of generality, we assume that all transmitted symbols  $\{s_k\}_{k=1}^K$  have equal variance  $\sigma_s^2 = P/K = 1$  such that  $\mathbf{R}_{ss} = \sigma_s^2 \mathbf{I}_K = \mathbf{I}_K$  and  $P = \text{Tr}(\mathbf{R}_{ss}) = K$ .

### 3.2 Channel Model

In order to capture the effect of the beam radiation patterns, we consider, without loss of generality, a geometry-based single-bounce scattering statistical model [12][13]. In the single-bounce approach, each transmit/receive path is broken into two subpaths: transmitter-to-scatterer and scatterer-to-receiver (described by their direction of departure, direction of arrival, and path distance). The scatterer itself is modeled typically via the introduction

of a random attenuation and phase shift. In this paper, we assume that the scatterers are randomly distributed on the surface of a sphere that is placed in the middle of the distance between the TXs and the RXs, as seen in Fig. 3. This model has been selected because, while it is relatively simple in comparison with other models, it captures small-scale fading and shadowing effects sufficiently well.

In order to facilitate performance analysis, we normalize the channel matrix. More specifically, for fixed  $\mathbf{H}$  (e.g., when a given channel realization over the coherence time of the channel is considered), the normalized channel matrix  $\tilde{\mathbf{H}}$  is given by [14]:

$$\tilde{\mathbf{H}} = a\mathbf{H}, \quad (6)$$

where, in the approach that we followed, the normalization constant is expressed as

$$|a|^2 = \frac{N_T N_R}{\|\mathbf{H}\|_F^2} = \frac{K^2}{\|\mathbf{H}\|_F^2}, \quad (7)$$

such that the power gain of the channel is

$$\text{Tr}(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger) = \|\tilde{\mathbf{H}}\|_F^2 = N_T N_R = K^2. \quad (8)$$

Of course, when the channel coefficients are RVs (e.g., longer time intervals are of interest), the expectation over  $\mathbf{H}$  should be taken in the appropriate expressions.

Note that in Eqs. (1)–(5), as well as throughout the paper, we denote the normalized channel matrix simply as  $\mathbf{H}$ , for convenience.

## 4 Radio Transmission Protocol

System operation is divided into three phases, namely, the learning phase, the selection phase, and the transmission phase. In the learning phase, each

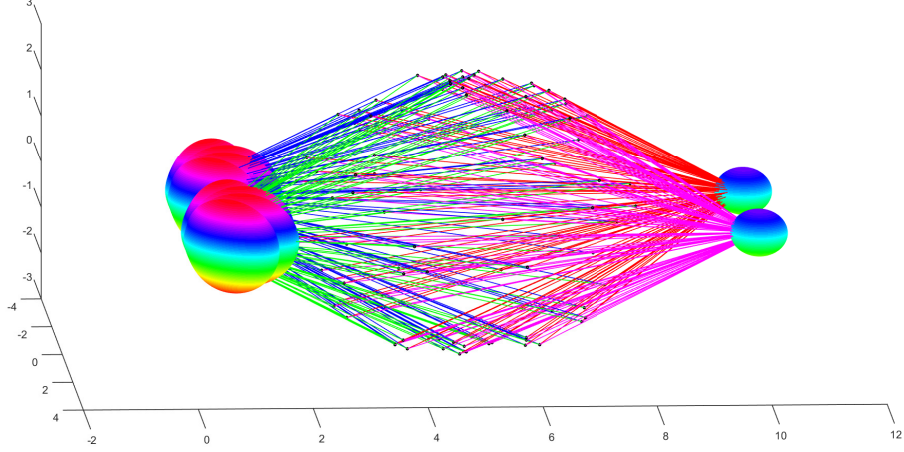


Figure 3: Single-bounce scattering model with 100 scatterers for a system with  $K = 2$  and  $L = 4$  beams per transmitter.

UE acquires information about the quality or the coefficients of each channel formed by the  $L^K$  different beam combinations at the TXs. In the selection phase, the BSs select jointly the best  $K$ -tuple of beams, according to the information that has been fed back to them by the UEs. Finally, in the transmission phase, communication takes place. The transmission strategy depends on the channel information available at the TXs.

#### 4.1 Learning Phase

There exist two different types of channel information that a UE may send back to the corresponding BS:

1. Its signal to interference plus noise ratio (SINR) for each beam combination at the TXs.
2. The CSI for the resulting “beam-domain transformed channel”, i.e., the complex-valued coefficients for the self- and cross-channels of the total  $K \times K$  channel.

#### 4.2 Beam Selection Phase

According to the type of channel information that the BSs acquired through feedback, the following beam selection criteria apply:

**Beam Selection Rule #1:** Select the  $K$ -tuple of beams that results in the largest sum of SINRs, i.e., the beam combination that maximizes  $\sum_{k=1}^K \gamma_k$ .

**Beam Selection Rule #2:** Select the  $K$ -tuple of beams that results in the equivalent channel matrix with the largest product  $\mathbf{H}\mathbf{H}^\dagger$ .

We should note that, even though our goal in this work is to generate beams using parasitic antenna arrays, since ESPAR technology reduces the cost, size, and complexity of the system, the general principle holds also for beams generated by conventional antenna arrays.

#### 4.3 Transmission Phase

In this phase, the transmission strategy that will be followed to send information symbols over the selected beams is determined. The performance metric that we wish to increase as much as possible while maintaining complexity and CSI feedback overhead low is the sum-rate throughput.

If beam selection is based on the measured SINRs at the RXs (i.e., the beams that maximize the sum of the received SINRs have been selected), then there is no additional processing taking place at the TXs. The sum-rate in this non-precoded transmission case is given by

$$R_{\text{NP}} = \sum_{k=1}^K R_k^{(\text{NP})} = \sum_{k=1}^K \log_2 \left( 1 + \gamma_k^{(\text{NP})} \right), \quad (9)$$

where

$$R_k^{(\text{NP})} = \log_2 \left( 1 + \gamma_k^{(\text{NP})} \right) \quad (10)$$

is the rate achieved by user  $k$  and  $\gamma_k$  is the SINR of user  $k$  which is expressed as

$$\begin{aligned} \gamma_k^{(\text{NP})} &= \frac{|h_{k,k}|^2 \sigma_k^2}{\sum_{m \neq k} |h_{k,m}|^2 \sigma_m^2 + \sigma_n^2} \\ &= \frac{|h_{k,k}|^2}{\sum_{m \neq k} |h_{k,m}|^2 + \sigma_n^2}. \end{aligned} \quad (11)$$

since  $\{\sigma_k^2\}_{k=1}^K = \sigma_s^2 = 1$ .

If, on the other hand, the channel matrices have been fed back to the TXs (i.e., the beams that result

in an equivalent channel with the largest product  $\mathbf{H}\mathbf{H}^\dagger$  have been selected), then the TXs jointly precode the transmit vector  $\mathbf{s}$ , thus transforming the MU-MISO IFC into a MU-MISO broadcast channel (BC).

We assume that linear precoding is employed by the system, in order to avoid the high RX complexity required by more advanced transmission techniques [13]. Linear precoding includes a family of simple but sub-optimal pre-processing techniques that exploit CSI at the transmitter (CSIT) to improve MU-MIMO performance—e.g., to increase the sum-rate throughput or to minimize the aggregated bit error rate. (Unless stated otherwise, perfect CSIT is assumed.)

A linear precoder generates a precoded signal vector as a linear transformation of the original symbol vector:

$$\mathbf{s}' = \mathbf{W}\mathbf{s}. \quad (12)$$

Hence, the received signal vector is expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{s}' + \mathbf{n}. \quad (13)$$

That is,

$$\underset{[K \times 1]}{\mathbf{y}} = \underset{[K \times K]}{\mathbf{H}} \underset{[K \times K]}{\mathbf{W}} \underset{[K \times 1]}{\mathbf{s}} + \underset{[K \times 1]}{\mathbf{n}}. \quad (14)$$

Equivalently, we can write

$$y_k = \mathbf{h}_k^\dagger \mathbf{w}_k s_k + \sum_{\substack{m=1 \\ m \neq k}}^K \mathbf{h}_k^\dagger \mathbf{w}_m s_m + n_k, \quad (15)$$

where  $\mathbf{h}_k \in \mathbb{C}^{K \times 1}$ ,  $\mathbf{w}_k \in \mathbb{C}^{K \times 1}$  and  $s_k$  are the channel vector, precoding vector, and data stream of user  $k$ , respectively;  $\mathbf{H}$  is the channel matrix; and  $\mathbf{W}$  is the precoding matrix expressed as

$$\mathbf{W} = [\mathbf{w}_1 \quad \cdots \quad \mathbf{w}_K]. \quad (16)$$

Note that the second term at the right-hand side of Eq. (15) represents the multi-user interference (MUI).

The SINR of user  $k$  for this linear precoding scheme is given by

$$\gamma_k^{(\text{LP})} = \frac{\|\mathbf{h}_k^\dagger \mathbf{w}_k\|^2}{\sum_{m \neq k} \|\mathbf{h}_k^\dagger \mathbf{w}_m\|^2 + \sigma_n^2}. \quad (17)$$

The expressions of the  $k$ th user's rate  $R_k^{(\text{LP})}$  and the sum-rate of the system  $R_{\text{LP}}$  are similar with the ones given in Eq. (10) and (9), respectively, for the non-precoded system.

The Tx power constraint is

$$\begin{aligned} \text{Tr} \left( \mathbb{E} \left\{ \mathbf{s}' (\mathbf{s}')^\dagger \right\} \right) &= \text{Tr} \left( \mathbb{E} \left\{ (\mathbf{W}\mathbf{s}) (\mathbf{W}\mathbf{s})^\dagger \right\} \right) \\ &= \text{Tr} \left( \mathbb{E} \left\{ \mathbf{W}\mathbf{s}\mathbf{s}^\dagger \mathbf{W}^\dagger \right\} \right) \\ &= \text{Tr} \left( \mathbf{W} \mathbb{E} \left\{ \mathbf{s}\mathbf{s}^\dagger \right\} \mathbf{W}^\dagger \right) \\ &= \text{Tr} \left( \mathbf{W} \mathbf{R}_{ss} \mathbf{W}^\dagger \right) \leq P. \end{aligned} \quad (18)$$

We propose the use of zero-forcing (ZF) precoding, which provides a promising compromise between complexity and performance [15], especially at the high signal to noise ratio (SNR) regime [16].

The reasoning behind ZF precoding is to employ a linear transformation to the transmit signal vector, according to Eq. (12), such that MUI becomes null, i.e., so that each user  $k = 1, \dots, K$  receives no interference from the signals intended for the other users [13]:

$$\|\mathbf{h}_k^\dagger \mathbf{w}_m^{(\text{ZF})}\| = 0, \quad m \neq k \quad (19)$$

Interference suppression at the TX is important in this setup, since single-antenna, non-cooperating RXs are unable to eliminate interference—the best they can do is to treat it as noise.

A common approach to accomplish this goal involves the inversion of the channel matrix at the TX in order to create orthogonal channels between the TX and the RXs (i.e., to diagonalize the effective channel). Channel inversion implies, under the flat-fading assumption, the multiplication of the original transmit vector signal with the right Moore-Penrose pseudo-inverse of the channel matrix  $\mathbf{H}$ , such that the matrix of the composite channel is the identity matrix. That is,

$$\mathbf{W}_{\text{ZF}} = \mathbf{H}^+ \Rightarrow \begin{aligned} \mathbf{s}' &= \mathbf{W}_{\text{ZF}} \mathbf{s} = \mathbf{H}^+ \mathbf{s} \\ \mathbf{H}\mathbf{W}_{\text{ZF}} &= \mathbf{H}\mathbf{H}^+ = \mathbf{I} \end{aligned} \quad (20)$$

where

$$\mathbf{H}^+ = \mathbf{H}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1}. \quad (21)$$

In practice, though, we typically normalize the precoding matrix  $\mathbf{W}_{\text{ZF}}$  in order to set the transmit power (after precoding) to a fixed value, independent of the channel  $\mathbf{H}$ , according to the given power constraint. An often used normalization, which is referred to as equal receive power (ERP) normalization, takes the following form [17]:

$$\mathbf{W}_{\text{ERP}} = \sqrt{\beta} \mathbf{H}^+ = \sqrt{\beta} \mathbf{H}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1}. \quad (22)$$

Under our assumption  $\sigma_s^2 = 1 \Rightarrow \mathbb{E} \{ \mathbf{s}\mathbf{s}^\dagger \} = \mathbf{R}_{ss} = \mathbf{I}_K \Rightarrow \text{Tr}(\mathbf{R}_{ss}) = K$  and by forcing the total transmit power after precoding to remain equal to  $P = K$ , that is,

$$\text{Tr} \left( \mathbf{W}_{\text{ERP}} \mathbf{R}_{ss} \mathbf{W}_{\text{ERP}}^\dagger \right) = K, \quad (23)$$

the power normalization factor  $\beta$  is given by [18]

$$\beta = \frac{1}{\text{Tr} \left( (\mathbf{H}\mathbf{H}^\dagger)^{-1} \right)} \quad (24a)$$

$$= \frac{1}{\left\| (\mathbf{H}\mathbf{H}^\dagger)^{-1} \right\|_F^2} \quad (24b)$$

$$= \frac{1}{\sum_{k=1}^K 1/(\lambda_k^2)} \quad (24c)$$

where  $\lambda_k$  is the  $k$ th singular value of  $\mathbf{H}$ .

When the channel matrix is square, as it is in our case where  $N_T = K$ , then the precoding matrix is simply given by [19]:

$$\mathbf{W}_{\text{ERP}} = \sqrt{\beta} \mathbf{H}^{-1}. \quad (25)$$

In any case, the corresponding received signal vector is given by:

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{s}' + \mathbf{n} \\ &= \mathbf{H}(\mathbf{W}_{\text{ERP}}\mathbf{s}) + \mathbf{n} \\ &= (\mathbf{H}\mathbf{W}_{\text{ERP}})\mathbf{s} + \mathbf{n} \\ &= \sqrt{\beta}\mathbf{s} + \mathbf{n}. \end{aligned} \quad (26)$$

Note that the scaling factor in Eq. (26) may lead in reduced SNR at the receiver if the channel matrix is ill-conditioned, i.e., if one of its singular values is very large.

The SINR at user  $k$  is given by Eq. (17) by setting the MUI at the denominator equal to zero and noting that  $\mathbf{H}\mathbf{W}_{\text{ERP}} = \sqrt{\beta}\mathbf{I}$ :

$$\gamma_k^{(\text{ERP})} = \frac{\beta}{\sigma_n^2}. \quad (27)$$

Therefore, all MUs experience the same SNR as expected, since the scaling factor is the same for all transmitted signals, and achieve the same rate

$$R_k^{(\text{ERP})} = \log_2 \left( 1 + \gamma_k^{(\text{ERP})} \right). \quad (28)$$

Hence, the sum-rate of this ZF precoding scheme is given by:

$$\begin{aligned} R_{\text{ERP}} &= \sum_{k=1}^K R_k^{(\text{ERP})} \\ &= \sum_{k=1}^K \log_2 \left( 1 + \gamma_k^{(\text{ERP})} \right) \\ &= \sum_{k=1}^K \log_2 \left( 1 + \frac{\beta}{\sigma_n^2} \right) \\ &= K \log_2 \left( 1 + \frac{\beta}{\sigma_n^2} \right) \end{aligned} \quad (29a)$$

$$= K \log_2 \left( 1 + \frac{1}{\sigma_n^2 \sum_{k=1}^K 1/(\lambda_k^2)} \right). \quad (29b)$$

Another commonly used normalization, which results in higher sum-rate than ERP normalization, is the so-called equal transmit power (ETP) normalization which is obtained by setting

$$\mathbf{F} = \mathbf{H}^+ = \mathbf{H}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1} \quad (30)$$

(or  $\mathbf{F} = \mathbf{H}^{-1}$  in case of a square channel matrix) and then dividing the elements of column  $k$  of  $\mathbf{F}$

with the norm of the corresponding column vector [20]:

$$\mathbf{W}_{\text{ETP}} = \frac{\mathbf{F}(:,k)}{\|\mathbf{F}(:,k)\|}, \quad k = 1, \dots, K. \quad (31)$$

The SINR in this case is

$$\gamma_k^{(\text{ETP})} = \frac{1}{\sigma_n^2 \|\mathbf{F}(:,k)\|^2} \quad (32)$$

and the sum-rate is given as

$$\begin{aligned} R_{\text{ETP}} &= \sum_{k=1}^K \log_2 \left( 1 + \gamma_k^{(\text{ETP})} \right) \\ &= \sum_{k=1}^K \log_2 \left( 1 + \frac{1}{\sigma_n^2 \|\mathbf{F}(:,k)\|^2} \right). \end{aligned} \quad (33)$$

Both these normalization methods set MUI to zero and have the same total transmit power. These ZF precoding design schemes aim at minimizing transmit power. Another design approach is to set

$$\mathbf{W}_{\text{WF}} = \mathbf{H}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1} \text{diag} \left( \sqrt{P_1}, \dots, \sqrt{P_K} \right) \quad (34)$$

and perform water-filling (WF) to optimally allocate transmit power per antenna  $\{P_k\}_{k=1}^K$  such that the sum-rate is maximized, under the total power constraint

$$\sum_{k=1}^K \|\mathbf{w}_k\| P_k = P. \quad (35)$$

However, this optimal ZF method is not applicable in the considered system setup due to the fact that it does not only increase the complexity of the cooperation between the BSs, but it requires also cooperation between the UEs (i.e., joint decoding).

## 5 Performance Results

In this section we evaluate the sum-rate performance of the proposed radio protocol via numerical simulations for the system setup described in Section 3 with  $K = 2$  and  $L = 4$ .

The performance results in Fig. 4 represent ergodic sum-rates achieved over a range of target SNR values at the receiver, from 0dB up to 30dB. These results have been obtained after 1,000 simulation runs by taking the expectation of the corresponding sum-rate equations. Also, in each simulation run, 100 sub-runs are used for the normalization of the channel matrices, as described in Section 3.2.

We consider transmission over the “beam-domain” channel with and without ZF precoding. That is, we study non-precoded transmission over the beams (which have been selected according to the SINRs that the MTs have fed back to the BSs)

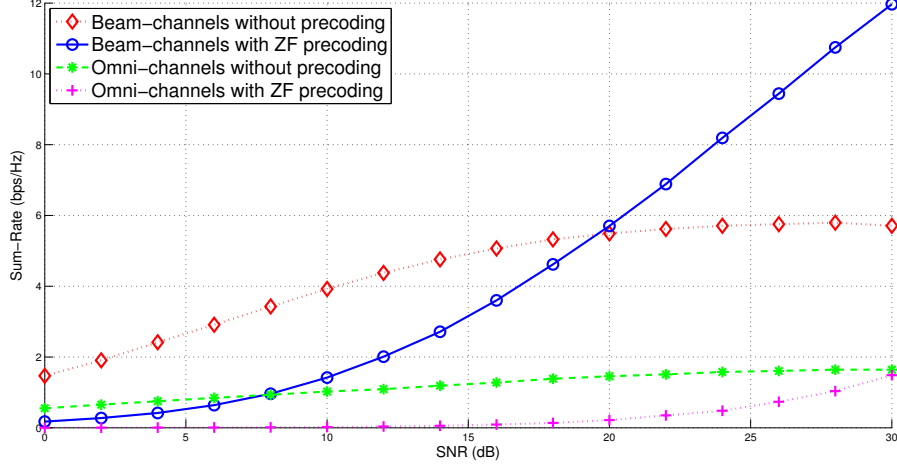


Figure 4: Sum-rate of different transmission schemes for the case of  $K = 2$  interfering links.

as well as the case where the MTs feed back the complex coefficients of the equivalent  $2 \times 2$  “beam-domain” channel and ZF precoding is incorporated in the transmission process.

For comparison purposes, we illustrate also the performance of the system when communication does not take place over a selected pair of beams, but we have instead conventional (non-precoded or precoded) “omni-channel” transmission. Of course, this type of non-precoded transmission would not take place in practice, since it is not a multi-user communication scheme, i.e., it does not take into account MUI; it is just included in the simulation as a lower performance bound.

In our simulations we have set, for mathematical tractability and without loss of generality,  $P = K$ .

We note that non-precoded transmission over the “beam-channels” outperforms significantly “omni-channel” transmission methods, even when the latter employ ZF precoding, especially at high SNR values. This is due to: (a) The gain that is introduced to the system as a result of the use of beams. (b) The fact that the TXs are informed about the SINR at the RXs and select the best beam combination each time.

We also note that the knowledge of the CSI at the transmitter for the equivalent “beam-domain transformed” channel and the joint precoding of the transmit vector that is performed based on that CSIT further improves the performance of the system at the high SNR regime.

Finally, it is worth noting that the sum-rate of the transmission schemes that incorporate ZF precoding keeps increasing linearly within the considered SNR range, whereas the non-precoded techniques experience an expected flooring of their performance due to the residual interference that they incur, which is not accounted for.

## 6 Complexity-Overhead and Complexity-Performance Tradeoffs

In Section 2, a brief overview of ESPAR and C-RAN technologies was given. In this Section, we present the complexity reduction accomplished by the use of the proposed radio transmission scheme described in Section 4, mainly in terms of channel estimation / channel information feedback, and the relevant complexity-overhead as well as complexity-performance tradeoffs. But first, let us summarize the main points of Section 2:

- ESPAR antennas reduce the size, complexity and cost of antenna arrays due to the fact that they use fewer active antenna elements and RF units than conventional antenna solutions and thus can be viewed as an enabler of massive MIMO.
- C-RAN facilitates the efficient cooperation between individual BSs in Coordinated Multi-Point (CoMP) setups since the centralization and virtualization of the BBUs as well as the use of optical fibers at the fronthaul for the connection of the BBUs with the RRHs is capable of meeting the stringent delay and bandwidth requirements that are imposed by this family of cooperative MIMO communication techniques.

The interested reader is encouraged to refer to the corresponding references cited in Section 2 for a more detailed presentation of the benefits and drawbacks / challenges of these technologies.

As we mentioned in Section 3, our focus is on the downlink transmission in a system setup with  $K$  cloud-based RRHs/BSs, each equipped with



a single-RF ESPAR, and  $K$  single-antenna UEs. Since the  $K$  RRHs can be viewed as a composite BS with  $K$  antennas due to the C-RAN architecture, MU-MISO communication techniques over a BC are applicable.

The use of ESPARs instead of conventional antenna arrays reduces significantly channel estimation complexity due to the decreased number of active antenna elements which is translated into a correspondingly decreased number of direct- and cross-channels. In the remaining Section, we will compare “omni-channel” and “beam-channel” transmission methods in terms of complexity assuming in both cases the use of ESPARs at the BSs, but we should keep in mind this important note, that is, that the replacement of antenna arrays by ESPARs already provides complexity reduction regarding channel estimation.

Initially, let us assume conventional transmission over “omni-channels”. The optimum precoding technique in terms of sum-rate is dirty paper coding (DPC) [13]. However, the computational complexity of this non-linear precoding scheme makes it impractical. ZF precoding, on the other hand, is a low-overhead scheme that completely eliminates MUI, but it amplifies noise power. Minimum mean square error (MMSE) considers noise in the precoding design process, leading to better performance than ZF in low SNR conditions, at the cost of increased complexity [13]. Hence, ZF precoding is the MU-MISO transmission method with the smaller overhead.

ZF precoding requires full CSIT. Since we have  $K$  active antennas at the composite TX (and recalling that we are considering narrowband channels), the equivalent channel is described by a  $K \times K$  matrix. In other words, each one of the  $K$  UEs should feed back to its corresponding BS a vector of size  $K$  containing the coefficient of the direct channel with that BS and the  $K - 1$  coefficients of the cross-channels with the other  $K - 1$  BSs.

If we assume non-precoded transmission, as a low complexity bound, then each UE should feed back to its BS only its SINR. Thus, the composite BS would have to collect  $K$  SINR values (real numbers), as opposed to the  $K^2$  channel coefficients (complex numbers) required in the ZF precoding scenario. Of course, as we mentioned in the previous Section, such a transmission scheme would not be used in practice due to its poor performance.

Now, let us consider the proposed radio protocol. Each one of the  $K$  RRHs is able to generate  $L$  distinct beams. Thus, there exist  $L^K$  possible beam combinations. Assuming non-precoded transmission (i.e., the beams to be used during transmission are selected according to criterion #1 described in Section 4), each UE should feed back to its BS its SINR for each one of the beam combinations. That

is, each one of the  $K$  UEs should send to its corresponding BS  $L^K$  SINR values. Therefore,  $K(L^K)$  real values will be sent back to the composite BS in total. In contrast, when ZF precoding is used over “omni-channels”,  $K^2$  complex values have to be sent back to the composite BS. However, in the latter case, accurate channel estimation is challenging when the number of BSs (and active TX antenna elements) is high, while the measurement of the received SINR value at each UE is a rather trivial task. Also, while non-precoded, SINR-based transmission over “omni-channels” presents the lowest overhead (the feedback of only  $K$  real values is required), it is not suitable for multi-user communication as we have already mentioned.

Let us turn our attention into ZF precoding-based transmission over the “beam-channels” (i.e., beam selection criterion #2 is applied). For each beam-combination, the required CSIT is represented by a  $K \times K$  complex-valued matrix. Thus, in total  $K^2(L^K)$  complex values have to be fed back to the composite BS instead of only  $K^2$  complex values that are required in the corresponding “omni-channels” case. However, a performance-complexity tradeoff is expected due to the “beam-gains” and the selection of the optimal “composite beam-domain channel”, as we have seen in Section 5.

Let us summarize: Non-precoded and precoded transmission over “beam-channels” results in an increase of the feedback overhead by  $L^K$  in comparison with the corresponding “omni-channel” schemes due to the fact that each UE should send back to its BS its SINR or channel matrix *for each beam combination*. However, it outperforms transmission over “omni-channels” and it simplifies (or even eliminates the need for) channel estimation. More specifically:

- “Beam-channel” communication techniques present a significant performance gain against their “omni-channel” counterparts.
- Even non-precoded “beam-channel” transmission outperforms ZF “omni-channel” transmission. Thus, by using predetermined beams, we could simply collect SINR values instead of estimating channel matrices, which is a much simpler procedure - and therefore, it is easier to be used in practice, even in large setups.<sup>1</sup>
- The increase in the channel information overhead caused by the use of beams can be partly compensated by the use of ESPARs instead of conventional antenna arrays, depending on

<sup>1</sup>In high mobility environments, it is possible to update the SINR value corresponding to each beam combination every time it is used to serve a user. Then, a record of time-windowed SINR values can be kept, thus further reducing the feedback requirements [21][22][23].

the ratio “reduction in the number of active antenna elements / increase in the number of beams”.

## 7 ZF Precoding-based Transmission over “Beam-Channels” with Imperfect CSIT

So far, we have assumed perfect CSIT. In practice, though, various sources of error (e.g., channel estimation errors, channel quantization errors, feedback errors etc.) may result in imperfect CSIT.

In this case, the imperfect channel matrix that is fed back to the TX is expressed as

$$\mathbf{H}_e = \mathbf{H} + \mathbf{E}, \quad (36)$$

where  $\mathbf{H}$  is the actual channel matrix and  $\mathbf{E}$  is an additive error matrix whose entries are independent and identically distributed (i.i.d.) RVs that are independent of  $\mathbf{H}$  and follow a  $[\mathbf{E}_{ij}] \sim \mathcal{CN}(0, \sigma_e^2)$  distribution.

Then, the precoded signal vector is expressed as

$$\mathbf{s}' = \sqrt{\beta} \mathbf{W}_{\text{ZF}} \mathbf{H}_e^+, \quad (37)$$

where the power normalization factor  $\beta$  is given by

$$\beta = \frac{1}{\text{Tr} \left( \left( \mathbf{H}_e \mathbf{H}_e^\dagger \right)^{-1} \right)} \quad (38a)$$

$$= \frac{1}{\left\| \left( \mathbf{H}_e \mathbf{H}_e^\dagger \right)^{-1} \right\|_F^2} \quad (38b)$$

$$= \frac{1}{\sum_{k=1}^K 1/(\lambda_k^2)}, \quad (38c)$$

and  $\lambda_k$  is the  $k$ th singular value of  $\mathbf{H}_e$ .

The received signal vector is expressed as

$$\begin{aligned} \mathbf{y} &= \mathbf{H} \mathbf{s}' + \mathbf{n} \\ &= \sqrt{\beta} \mathbf{H} \mathbf{H}_e^\dagger \mathbf{s} + \mathbf{n} \\ &= \sqrt{\beta} (\mathbf{H}_e - \mathbf{E}) \mathbf{H}_e^\dagger \mathbf{s} + \mathbf{n} \\ &= \sqrt{\beta} \mathbf{s} - \sqrt{\beta} \mathbf{E} \mathbf{H}_e^\dagger \mathbf{s} + \mathbf{n}. \end{aligned} \quad (39)$$

The SINR of each user, given that  $\sigma_s^2 = 1$ , is

$$\gamma_k^{(\text{ZF})} = \frac{\beta}{P\sigma_e^2 + \sigma_n^2} = \frac{1}{(\sigma_e^2 + 1/P) \sum_{k=1}^K 1/\lambda_k^2}, \quad (40)$$

while the sum-rate is expressed as

$$R_{\text{ZF}} = K \log_2 \left( 1 + \frac{1}{(\sigma_e^2 + 1/P) \sum_{k=1}^K 1/\lambda_k^2} \right). \quad (41)$$

Note that the variance of the error,  $\sigma_e^2$ , limits the sum-rate, i.e., as  $P \rightarrow \infty$ ,

$$R_{\text{ZF}} = K \log_2 \left( 1 + \frac{1}{(\sigma_e^2) \sum_{k=1}^K 1/\lambda_k^2} \right). \quad (42)$$

We conclude this Section by evaluating the sum-rate performance of ZF precoding transmission over pre-determined beams for the considered setup in the case where CSIT is imperfect and for a range of error variance values  $P_e$  from  $10^{-3}$  up to  $10^{-1}$  via simulations and we compare these results with the performance obtained when CSIT was assumed to be perfect.

As we see in Fig. 5, when the error variance is small, the degradation in the sum-rate throughput is negligible and only noticeable at the high SNR regime. However, as the error variance increases, the sum-rate is decreased and at some point floors (the higher the error variance, the sooner this flooring takes place).

## 8 Conclusions

In this paper, we proposed a new radio transmission protocol for a system setup that could be an enabler of next-generation MU-MIMO systems. The proposed scheme performs extremely well in terms of the achieved sum-rate while it reduces significantly the complexity of the system. In the future, we plan to extend this work by studying larger setups and cases with imperfect CSI feedback.

## Acknowledgments

This work has been supported by the EC FP7 project HARP (<http://www.fp7-harp.eu/>) under grant number 318489.

## References

- [1] ITU-R, “Assessment of the global mobile deployments and forecasts for international mobile telecommunications,” Report M.2243, November 2011.
- [2] “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014-2019,” Cisco, White Paper, February 2015.
- [3] M. Sawahashi *et al.*, “Coordinated Multi-point Transmission/reception Techniques for LTE-Advanced [Coordinated and Distributed MIMO],” *IEEE Wireless Communications Magazine*, vol. 17, no. 3, pp. 26–34, 2010.

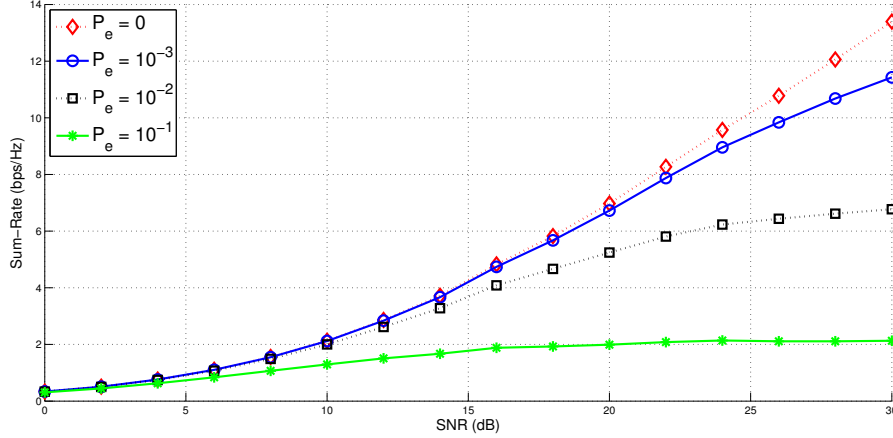


Figure 5: Sum-rate of ZF precoding-based transmission over “beam-channels” for the case of imperfect CSIT and for various values of error variance.

- [4] D. Lee *et al.*, “Coordinated Multipoint Transmission and Reception in LTE-Advanced: Deployment Scenarios and Operational Challenges,” *IEEE Communications Magazine*, vol. 50, no. 2, pp. 148–155, 2012.
- [5] Massive MIMO Project. [Online]. Available: [www.massivemimo.eu](http://www.massivemimo.eu)
- [6] T. L. Marzetta, “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.
- [7] A. Kalis, A. Kanatas, and C. B. Papadias, Eds., *Parasitic Antenna Arrays for Wireless MIMO Systems*. Springer-Verlag New York, 2014.
- [8] “C-RAN: The Road Towards Green RAN,” White Paper, China Mobile, Tech. Rep., October 2011.
- [9] HARP EU FP7 Project. [Online]. Available: <http://www.fp7-harp.eu/>
- [10] CPRI Protocol. [Online]. Available: <http://www.cpri.info/>
- [11] A. Checko *et al.*, “Cloud Ran for Mobile Networks - a Technology Overview,” *IEEE Communications Surveys and Tutorials*, vol. 17, no. 1, pp. 405–426, 2014.
- [12] P. Almers *et al.*, “Survey of Channel and Radio Propagation Models for Wireless MIMO Systems,” *EURASIP Journal on Wireless Communications and Networking*, 2007.
- [13] H. Huang, C. B. Papadias, and S. Venkatesan, *MIMO Communication for Cellular Networks*. Springer US, 2012.
- [14] S. Louka and G. Levin, “On Physically-based Normalization of MIMO Channel Matrices,” *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, 2009.
- [15] A. H. Mehana and A. Nosratinia, “Diversity of MIMO Linear Precoding,” *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 1019 – 1038, February 2012.
- [16] T. Yoo and A. Goldsmith, “On the Optimality of Multiantenna Broadcast Scheduling using Zero-Forcing Beamforming,” *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, March 2006.
- [17] Y. Cho *et al.*, *MIMO-OFDM Wireless Communications with MATLAB*. Wiley-IEEE Press, 2010.
- [18] M. Joham, W. Utschick, and J. Nosssek, “Linear Transmit Processing in MIMO Communications Systems,” *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2700–2712, August 2005.
- [19] C. B. Peel *et al.*, “A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication-Part I: Channel Inversion and Regularization,” *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, January 2005.
- [20] M. Ghosh, “A Comparison of Normalizations for ZF Precoded MU-MIMO Systems in Multipath Fading Channels,” *IEEE Wireless Communications Letters*, vol. 2, no. 5, pp. 515–518, July 2013.
- [21] P. Viswanath, D. N. C. Tse, and R. Laroia, “Opportunistic beamforming using dumb antennas,” *IEEE Transactions on Information*

*Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

- [22] D. Avidor, J. Ling, and C. B. Papadidas, “Jointly opportunistic beamforming and scheduling (JOBS) for downlink packet access,” in *IEEE ICC*, vol. 5, 2004, pp. 2959–2964, Paris, France.
- [23] K. Ntougias, N. Taramas, and C. B. Papadidas, “Low-feedback Cooperative Opportunistic Transmission for Dynamic Licensed Shared Access,” in *EUSIPCO*, 2015, to appear.